Landscape and Urban Planning 191 (2019) 103436

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Research Paper

Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing

Jingxian Tang, Ying Long*

School of Architecture and Hang Lung Center for Real Estate, Tsinghua University, Beijing, China

ARTICLE INFO

Keywords:
Public space
Visual quality
Perceived quality
Street view picture
Hutong

ABSTRACT

Although it is widely known that quality of street space plays a vital role in promoting urban vibrancy, there is still no consensus on how to quantitatively measure it for a large scale. Recent emerging dataset Street View Picture has revealed the possibility to overcome the previous limit, thus bringing forward a research paradigm shift. Taking this advantage, this paper explores a new approach for visual quality evaluation and variation identification of street space for a large area. Hutongs, which typically represent for historical street space in Beijing, are selected for empirical study. In the experimental part, we capture multi-years Tencent Street View Picture covering all the Hutongs, and conduct both physical and perceived visual quality evaluation. The physical visual quality of street space is achieved automatically by combining 3-dimensional composition calculation of greenery, openness, enclosure using machine-learning segmentation method SegNet, and 2-dimensional analysis of street wall continuity and cross-sectional proportion; perceived visual quality of street space is evaluated by stay willingness scoring from five aspects. The variation of quality is evaluated based on the identified physical space variations. The result indicates that visual quality of Hutongs are not satisfied, while some regeneration projects in the historical protection block is better. Most Hutongs are in shortage of visual green, relative more continuous but with low cross-sectional ratio. Hutongs near main road witness an increasing trend of motorization. The difference between physical and perceived quality indicates the feasibility and limitation of the auto-calculation method. In the most recent 3–4 years, less than 2.5% Hutongs are improved, which are mainly slow beautification.

1. Introduction

Streets, together with parks, plazas and squares, have been widely deemed as public space, which is an important arena for urban vitality. Streets, as a vital component of the built environment, are highlighted for their values across various dimensions, such as traffic, aesthetics, public health and neighborhood interaction (Jiang, Chang, & Sullivan, 2014; Mounir, Attia, & Tayel, 2016; Middleton, 2016). Good street environments could enhance the frequency of outdoor activities, thus affecting the users' behavior. Fuelled by the increasing emphasis on **Quality of Street Space** (termed "**QoSS**" in this paper) and lively built environment, designing human-oriented street space has become one of the crucial environmental improvement strategies for competitive cities. Various initiatives have been proposed for streets, such as *Better Streets* in London, *Green Streets* in the city of Portland, *Cycle Strategy* in Copenhagen, and *Street Design Guidelines* in Shanghai. With this background, recent years have seen the emergence of several street environment evaluation platforms like *Walk Score, State of Place Index, Ratemystreet,* and *Walkability Asia*. The widespread street policies and platforms indicate a global consensus on highlighting QoSS in addition to extensive discussion on the traffic function of streets.

This call for the renewed livelihood of streets is not the first attempt historically. Efforts trace back to the 1960s when Jacobs (1961) used social descriptions to express the pursuit of neighborhood life. Other scholars also interpreted it into design vocabularies to define the concept of good street space (Lynch, 1960; Ewing & Clemente, 2013). However, due to the time-consuming nature of field surveys, nearly all urban designers and researchers at the time were restricted to small-scale empirical studies, finding connections between physical appearance and social attributes and deepening their logical reasoning through induction and deduction, instead of principle exploration based on large-scale investigation. Automatic quantitative measurement of spatial QoSS for large areas has long faced difficulties, and problems compound with temporal variation. This situation prevents planners and designers from scientifically proposing a rational blueprint for streets.

---

* Corresponding author.
E-mail address: ylong@tsinghua.edu.cn (Y. Long).

https://doi.org/10.1016/j.landurbplan.2018.09.015
Received 14 July 2017; Received in revised form 29 August 2018; Accepted 16 September 2018

The emerging dataset Street View Picture (SVP) and cutting-edge techniques like deep learning algorithms for processing images shed light on measuring the visual aspect of QoSS and its variation over a large geographical area. This study builds a methodology to measure the **visual Quality of Street Space** (vQoSS) as well as its temporal variations; the approach is then applied to historical streets in Beijing (Hutongs). The research questions are as follows: (1) What is the best way to measure the physical visual quality of street space (termed "**physical vQoSS**") automatically on a large scale? (2) What is the overall perceived visual quality of street space (termed "**perceived vQoSS**") in the study area, and to what extent can the physical vQoSS represent the perceived one? (3) What is the best method for detecting the physical changes of street components so as to measure the **vQoSS** variation over time?

This paper reviews the related definitions, theories, and approaches regarding quantitative measurement of vQoSS and its variation in Section 2. This is followed by an introduction of the study area and data collection, as well as details about the research methods in Sections 3 and 4, respectively. In Section 5, an empirical study of Hutongs has been conducted to test and explore the feasibility and limitations of the methods. The paper also reflects upon the status quo and the changing trends of temporary design for the historical area in Beijing. The evaluation result of Hutongs and the visual quality changes in recent years are examined. The last section gives the concluding remarks, and proposes potential applications and limitations of the study.

2. Literature review

2.1. Defining "visual quality of street space" (vQoSS)

In this paper, the concept of "**quality**" is used to address multi-aspect attributes of street space. Quality generally refers to a degree of excellence in something, and it is applied to the concept of street spaces as it relates to their environmental conditions and level of service. Fruin (1971) emphasizes the synergy of safety, security, convenience, continuity, comfort, and attractiveness of streets in general; Slater (1985) proposed the concept of comfort and classified street comfort into three categories, namely physical, psychological, and physiological. Physiological comfort is perceived through different sensors such as color, light, smell, sound and thermal comfort, which are based on the perception of the physical compositions of street space. To be more specific, a street equipped with appropriate form and perceptual design may gain increased physical function and emotional appeal, thus having restorative potential as mentioned previously by Kaplan (Kaplan & Kaplan, 1989; Kaplan, 1995). In addition, the social meaning or symbolic image of a space can be attributed to its perceived feeling in general (Lynch, 1960; Thompson, 2002; Fyfe, 2006). The overall quality generally invokes a combination of visual and non-visual feelings, influenced by weather, climate, pollution, human activities, as well as the context attached to certain history (Nasar, 1989; Cassidy, 1997; Johansson, Sternudd, & Ferreira, 2015).

During the last ten years, a considerable number of studies have measured the multisensory qualities of street space, including their auditory quality (Pheasant, Horoshenkov, Watts, & Barret, 2008; Easteal, Bannister, & Kang, 2014; Aletta, Lepore, Kostara-Konstantinou, Kang, & Astolfi, 2016; Aiello, Schifanella, Quercia, & Aletta, 2016), the smellscape (Henshaw, 2014), and the visual streetscape (Li, Ratti, & Seiferling, 2017; Harvey, Aultman-Hall, Hurley, & Troy, 2015; Naik, Kominers, Raskar, Glaeser, & Hidalgo, 2015; Salesses, Schechtner, & Hidalgo, 2013). In this paper, the focus is the development of methodology for measuring the visual streetscape. vQoSS is the comprehensive environmental condition and users' physical reflection of the static image of the space, including its color, texture, scale, and styles. On the one hand, vQoSS can be inferred from users' willingness to remain in the space, while on the other, it objectively relies on physical components of streetscape, namely tree canopies (greenery), enclosures, openness, connectivity, street wall continuity, density, accessibility, cross-sectional proportion, scale, tidiness and so on (Harvey, Aultman-Hall, Hurley, & Troy, 2015; Alexander et al., 1977; Saelens, Sallis, & Frank, 2003).

Classic academic works have repeatedly proven that physical space and street form are the cornerstones of street activities (Lynch, 1984; Gehl & Svarre, 2013; Gehl, 2013). From the perspective of environmental psychology, they play a fundamental role in social interaction, while good form brings better perceived visual feeling. In sum, greenery, openness, enclosures, street wall continuity, and cross-sectional proportion, are some of the basic and over-discussed morphological features among all the variables that contribute to the vQoSS.

Enclosure forms space; it represents the extent of human-scale; a well enclosed street tends to produce an impression of higher security for users, thus providing more opportunity for physical activities (Arnold, 1980; Alexander et al., 1977; Owens, 1993; Rapoport, 2016; Ewing & Handy, 2009), while a wide set-back structure always generates a feeling of emptiness and inactivity (Dover & Massengale, 2013; Montgomery, 2013). Therefore, enclosure is highlighted as the primary feature in designs that amplify street life and it is widely used as a measurement of comfort. **Continuity** could be defined as the proportion of street edge intersecting with buildings (Harvey, 2014; Heath et al., 2006); a complete and continuous building façade generates a streetscape with order and vitality within the provided enclosure (Ewing & Handy, 2009; Gehl & Svarre, 2013). With commercial frontage, a continuous street will attract more pedestrians in its closed-off atmosphere (Gehl, 2013). The contribution of **greenery** to vQoSS and perceived safety has been repeatedly documented (Li, Zhang, & Li, 2015, Li, Zhang, Li, Ricard, et al., 2015). Jiang et al. (2015) proved that tree cover density influences peoples' stress, and a tree canopy augments the enclosed feeling.

Cross-sectional proportion affects pedestrians' perception of scale, influencing their willingness to remain in a space (Dover & Massengale, 2013; Ewing & Handy, 2009). A wider street always implies higher motorization with more traffic, and potential danger, pollution, noise and car accidents. By contrast, a relatively narrow street encourages public leisure (Gehl & Svarre, 2013; Gehl, 2013). Harvey et al. (2015) verify that the skeletal proportions of streetscapes across New York City accounted for approximately 42% of the variability in perceived safety, and both enclosing buildings and tree canopies had a positive impact on vQoSS. Last but not least, **openness**, the degree of sky visibility, determines the amount of perceived lightness, having an impact on visual perception and pleasantness (Bruce, Green, & Georgeson, 2003; Li et al., 2017), and microclimate (Lin, Tsai, Hwang, & Matzarakis, 2012; Grimmond, Potter, Zutter, & Souch, 2001; Unger, 2008). Besides these five aspects, there are many indicators that reflect the physical quality of the street. However, in this paper, a synthesis method is demonstrated against the currently measurable ones. Other variables could be added into the framework gradually in further empirical and theoretical studies.

2.2. Key methods for measuring vQoSS

Academic discussions and applications of quantitative measurement of vQoSS have revolved around five categories: (1) subjective perception assessment; (2) systematic observation and rating of streetscape; (3) physiological monitoring and analysis; (4) laboratory experiments and analysis; and (5) computer-assisted auditing and evaluation. In addition, some discussions integrate two of the above methods for one location. These are generally addressed in Table 1.

2.2.1. Subjective perception assessment

Subjective perception assessment is the method of evaluating respondents' perceived emotions to the selected built environment through different kinds of interviews or in-person questionnaires. The results could help to identify important factors in the built environment

Table 1
Classification of methods for measurement of vQoSS.

Category	Sub-Category	Sample	Research method
A. Subjective perception assessment	Interview	McGinn, Evenson, Herring, Huston, and Rodriguez (2007)	Perception of the built environment gathered from responses to 1270 telephone surveys.
	In-person questionnaires	Sallis, Johnson, Calfas, Caparosa, and Nichols (1997)	43 in-person questionnaires were used to assess environmental variables at homes, in neighborhoods, or on frequently traveled routes.
B. Systematic observation and rating of streetscape;	Based on site survey	Gehl and Gemzøe (1996)	Investigation on Public Life and Public Space (PLPS), measuring quality of public space, quality of the facade along the street, pedestrian volume, and activities.
	Based on video clip or videotaping	Ewing and Clemente (2013)	Construction of systematic evaluation index, including five aspects to represent the quality of the street, and the video recording of interviewees' subjective ratings of the street.
	Based on large -scale collected SVP	Tang et al. (2016)	Subjective rating of the user's willingness to remain in the street space using multi-year SVPs, to measure the comprehensive quality and temporal variation.
		Curtis, Curtis, Mapes, Szell, and Cinderich (2013)	Using Google Street View GSV as a systematic observation tool for built environment audits.
		Rundle, Bader, Richards, Neckerman, and Teitler (2011)	Evaluation of the aesthetics and other physical aspects of the neighborhood with SVPs.
C. Physiological monitoring and analysis	Electrophysiological evaluation by sensors	Jiang et al. (2014)	Measurement of the participants' stress in different eye-level tree cover densities by measurement of skin conductance and salivary cortisol levels.
	Observation of users with other wearable devices	Aspinall et al. (2013)	Observation of the differences of Electroencephalograms (EEG) characteristic of participants' response in different walking environments, with the help of Emotiv EPOC recorders.
	Eye-tracking Metrics	Dupont, Antrop, and Van Eetvelde (2014)	Using Eye-tracking Metrics (ETM) to investigate how people perceive and observe landscape images (not street quality, but similar) in the lab.
D. Laboratory experiments and analysis	Virtual reality	Luigi et al. (2015)	Using Immersive Virtual Reality (IVR) as a tool to stimulate the different human senses and create the experience of the sense of "being there".
	Augmented reality (AR)	Carozza et al. (2014)	A novel AR system based on monocular vision is built to calculate occlusions of the built environment on augmenting virtual objects.
E. Computer-assisted auditing and evaluation	GIS and remote sensing analysis based on 2D data	Kendal, Williams, and Williams (2012)	Using aerial photos with a 10 cm resolution as the data source and conducting GIS analysis of tree cover in gardens, parks, and streetscapes.
		Harvey (2014)	Measuring continuity and cross-sectional proportion of streetscape with GIS.
	Visual analysis or auditing tool of SVP	Bader et al. (2015)	Using GSV to conduct virtual audits of neighborhood environments with Computer Assisted Neighborhood Visual Assessment System (CANVAS).
		Li, Zhang, Li, Ricard, et al. (2015)	Green vegetation extraction from GSV and calculation of the green view index of the streetscape.
	Machine-learning-assisted method for large-scale SVP	Salesses et al. (2013)	Evaluation of street quality with regard to the aspects of perception of safety, class and uniqueness based on a dataset created by Place Pulse.
		Naik et al. (2014)	Using 'Streetscore', a scene-understanding algorithm to predict the perceived safety of a streetscape, using training data from an online survey with contributions from more than 7000 participants.
		Harvey et al. (2015)	Identification of a set of "streetscape skeleton" design variables, and measurement of these variables' effect on a large and diverse sample of streetscapes, examining their relationship to perceived safety scores created by Place Pulse.
		Tang and Long (2017)	Evaluation of quality with both subjective rates from five aspects created by Ewing and Clemente (2013), and objective component analysis using image segmentation by SegNet.
		Shen et al. (2018)	Integration of visualization techniques with machine learning models to facilitate the detection of street view patterns with StreetVizor.

that affect physical activities (Brownson, Hoehner, Day, Forsyth, & Sallis, 2009). For example, Appleyard and Lintell (1972) carried out field interviews on three similar streets with different traffic levels in San Francisco to find out how traffic conditions affected the livability and quality of the street environment. The measurement includes five indicators, namely perceived livability, levels of social interaction, territorial extent, environmental awareness and safety for street quality.

2.2.2. Systematic observation and rating of streetscape

Systematic observation and rating is the procedure to obtain a complete understanding of the physical features of streetscape. The data source for this method includes survey images, video clips (taping), and SVPs. Evaluation indexes with different indicators have been constructed to meet a baseline requirement to measure the otherwise immeasurable (Ewing & Clemente, 2013; Ye & van Nes, 2014). These indicators convert the blurry conception of vQoSS into understandable vocabularies related to urban design, facilitating designers' pre-assessment and shaping of design ideas. For example, Ewing and Clemente (2013) construct a systematic evaluation index including Enclosure, Human scale, Transparency, Tidiness and Imageability to represent street quality. Tang and Long (2017) decompose the vQoSS according to the above five dimensions and attach a subjective score from designers. Human observers allow audits to account for variables that are subjective and nuanced, but sometimes impractical to evaluate from existing spatial datasets (Harvey, 2014). It is a reliable and preferable method once we overcome the limitations in audit scale and raise efficiency with the help of new technology.

2.2.3. Physiological monitoring and analysis

Some researchers conduct technologically advanced evaluation with the help of biosensors to record the physiological response of pedestrian groups in different street spaces, for example using salivary cortisol measurement, Eye-tracking Metrics (ETM), Electroencephalograms (EEG), nuclear magnetic resonance (NMR), and hormone tests. These technologies provide a new perspective for street quality measurement. For example, Aspinall, Mavros, Coyne, and Roe (2013) investigate the use of mobile EEG to record the emotional experience of a group of walkers in three types of street space. Jiang (2015) measured participants' stress level changes in neighborhood street scenes through hormonal (salivary cortisol), physiological (skin conductance) and psychological (self-reported stress) variables. These devices enable the observation of specific physiological parameters of users, reflecting the differences between physical street environments.

2.2.4. Laboratory experiments and analysis

Several studies have emphasized the possibility of Virtual or augmented reality (VR and AR) as visual stimuli for subjective perception research in the laboratory (Wang, Kim, Love, & Kang, 2013). Combined with radio processing and 3D modeling, a simulated environment with controllable variables could be set in VR, enabling the researchers to build experiments to evaluate perception, which could be integrated in the urban design process (Luigi, Massimiliano, Aniello, Gennaro, & Virginia, 2015). AR and VR provide advantages for enhancing state-of-the-art visual scenes through proposed designs within the real environment (Carozza, Tingdahl, Bosché, & Gool, 2014).

2.2.5. Computer-assisted auditing and evaluation

Computer-assisted auditing and evaluation becomes increasingly popular due to efficiency. With GIS and remote sensing, aerial photos with high resolution are used for top view analysis. And the emergence of the Street View Picture dataset has brought more possibilities. SVPs create a continuous 360° image of a streetscape (Li, Zhang, & Li, 2015, Li, Zhang, Li, Ricard et al., 2015), allowing users to easily imagine the real vQoSS scene. It presents all surrounding information equivalent to an on-the-spot investigation on a larger scale at one time. In addition, companies such as Tencent even have a multi-year picture recording function to record a time series of variance in street space. In recent years, increasing numbers of quantitative methodologies and empirical investigation studies have been explored to use SVP data in existing methods. Harvey (2014) identifies four streetscape skeleton types of three American cities with the help of SVPs and GIS. Combined with machine-learning, MIT Media Lab launched a project called Place Pulse, aiming to quantitatively identify which areas of a city are perceived as wealthy, modern, safe, lively, active, central, adaptable or family friendly, through the online website Place Pulse (http://pulse.media.mit.edu) and large-scale SVPs. Further studies by Naik, Philipoom, Raskar, and Hidalgo (2014), Salesses, Schechtner, and Hidalgo (2013), Harvey, Aultman-Hall, Hurley, and Troy (2015), and Li, Zhang, and Li (2015) were based on the dataset of Place Pulse. The latest exciting method from SegNet presents a deep learning framework for semantic segmentation of SVPs through which 12 categories of objective elements are recognized (Kendall et al., 2015). Together with other similar methods, this breakthrough increases the potential for tremendous advances in the field of urban design, especially in studies of the built environment. Tang and Long (2017) firstly apply image segmentation in their comparative study of the street quality in central Beijing and Shanghai, illustrating its feasibility.

2.3. Key factors for identifying variations of vQoSS

The automatic quantification of spatial variation is not easy to achieve. Social scientists interested in gentrification apply community level statistical data to discover the interaction between spatial change and the social attributes instead of using real spatial appearance variables. Computational science brings a recent breakthrough. Both perceived and physical variation detection are shown in Table 2. Naik et al. (2015) firstly use the *Urban Change Coefficient*, calculating the dynamic perceived safety variation over a period of time derived from an algorithm checking SVPs. Using deep Convolutional Neural Networks and the applied algorithm, streetscape level variations in pictures and videos can be identified. Physical variation identification through scoring is also more accurate, and avoids the errors generated by view angles. Tang, Ma, Zhai, and Long (2016) divide the physical street space into four parts according to sections and identify the temporal variation, building an evaluation system on physical change and its effectiveness, which indicates visual quality differentiations from year to year. This method is relatively reliable and easy to carry out, and is therefore used in identification of Hutong variation to reflect the change of vQoSS in this study.

Based on the above understandings, this research proposes a comprehensive approach for evaluation and variation identification of vQoSS with general applicability, not only with the traditional subjective scoring by designers but also with the automated recognition of the physical visual quality to reduce subjective bias.

3. Methodology

3.1. Study area and data collection

3.1.1. Study area

Our research is conducted within the Second Ring Road of Beijing (with an area of 62 km^2), also called the Old City, which dates back to Yuan Dynasty. Fig. 1 presents the road map of the area. In the Ming and Qing Dynasty, the Old City was the outer boundary of Beijing, mainly consisting of main roads (*Da jie*), secondary roads (*Xiao jie*), alleys and the Chinese traditional courtyard (*siheyuan*). According to the *Record Draft on Blocks and Streets in Beijing* by Zhu Yixin in the Qing Dynasty, there were about 3500 alleys and 978 of them were directly called Hutongs. Nowadays, after years of dramatic revolution and modernization, not all alleys retain their original historical form. About 2073 alleys, accounting for 13% of all, survive under the historical block protection policy and keep their traditional status as Hutongs, these are highlighted in red in Fig. 1. As a footnote, this map is generated from the 2015 road/streets map of Beijing obtained from a local road navigation firm, and all of the Hutongs have been simplified into single-line streets using ArcGIS.

3.1.2. Hutongs

Hutongs are narrow pathways or alleys enclosed by two rows of traditional *siheyuan.* In the Yuan Dynasty, the city was divided into several rectangular residential blocks by intersecting main roads. In each residential block, narrow streets and Hutongs were interlaced to form the basic skeleton (Fig. 2). The width of a Hutong is 6 steps (traditional Chinese unit of length), about 9.3 m in metric units, serving as the corridors for residents' daily life in that time. The skeleton is small-scale and has few trees beside the wooden residential buildings to avoid fires. In a nutshell, Hutongs typically represent the fabric of historical streets and are of significant importance to Beijing. From the perspective of historical value, there is an urgency for a technical recognition and interpretation of the characteristics of traditional street forms. The restoration, regeneration and transformation of Hutong spaces are key to historical protection in Beijing, and need continuous scientific observation. While this paper emphasizes mostly the methodology, the completeness and pertinence of observations of Hutongs are relatively preliminary, and waiting for further research.

3.1.3. Street view pictures from Tencent

This study uses the SVP dataset from Tencent to investigate the visual quality of Hutongs. The Tencent Map (http://map.qq.com) was selected to make use of its 'Time Machine' function. A user can find the 'Time Machine' button once entering into the Tencent SVP browser, which is a database of all past images between 2012 and 2016. As a

Table 2
New methods for variation detection.

Category	Sample	Research method
A. Perceived variation detection based on algorithm	Naik et al. (2015)	Based on the measurement method of Streetscore published in Naik et al. (2014), calculation of *Urban Change Coefficient* of streets from 2007 to 2014.
B. Automatic physical variation detection	Taneja, Ballan, and Pollefeys (2011)	Using low-resolution images to detect geometric changes for urban environment with algorithm.
	Sakurada and Okatani (2015)	Detection of variation of streetscape from street image pair using Convolutional Neural Networks (CNNs), features and super-pixel segmentation
	Alcantarilla, Stent, Ros, Arroyo, and Gherardi (2016)	Detection of change in street video clip by deep CNNs.
C. Physical variation detection by scoring	Tang et al. (2016)	Division of the street space into four parts according to sections and identification of temporal variation by scoring.

Fig. 1. Map of Second Ring Road and Hutongs in Beijing.

departure from previous studies, data is captured directly from the Tencent website with open source navigation script processing, building a simulation browser rather than requesting through the Tencent SVP API, because historical SVPs could not be captured directly by the API.

Data acquisition is divided into three steps. First of all, we generate investigation points of Hutongs from the above single-line map at intervals of 50 m, each point with latitude and longitude parameters to facilitate the following scrawling process. Second, 'Panos' was obtained, the Tencent view identification, from geographical coordinates (Long, Liu, & Tang, 2016). Then with parameters including size, ID, pitch (pitch angle), heading (yaw angle) (In our research,: size: 1280 * 720; pitch: 0; heading: 0, 90, 180, 270 for four directional SVPs), and the request Key, we could get an URL to open the street view website pages.

The following is an URL example (for Pano sample in this link: 10011022120723095812200):

http://apis.map.qq.com/ws/streetview/v1/image?size = 600x480&pano = 10011022120723095812200&pitch = 0&heading = 0&key = OB4BZ-D4W3U-B7VVO-4PJWW-6TKDJ-WPB77.

Afterward, with the support of NodeJS, CasperJS and SlimerJS, we simulate the artificial crawling process from is simulated through entering into a Tencent SVP browser, clicking the Time machine button and taking an SVP screenshot. Through the above procedures, we snaps of multi-year SVPs of four directions, namely north, west, east and south, are obtained, the stitching of which forms a panoramic view. Finally, 51,356 images for 1892 locations are captured (accessed on May 18, 2016, some locations have multi-month data in one year).

3.1.4. Other data

GIS data of buildings are also used, each of which has a footprint in polygons and including building height. There are 213,688 buildings in the old city in total. The buildings are used for calculating streetscape parameters to feed the vQoSS evaluation model.

3.2. Methodology

Fuelled by the above questions, a method framework is built containing three main parts, which is shown in Fig. 3.

(1) Automatic evaluation of physical vQoSS is done by selecting the typical physical feature variables, exploring an auto-calculated approach with the support of 3-dimensional composition calculation, including machine-learning image segmentation technology, and 2-dimensional GIS analysis. Greenery, openness, enclosures, street wall continuity and cross-sectional proportion are taken as the representative elements; these five measurable aspects (Fig. 4) are selected according to the literature review and their capacity as measurable aspects, to demonstrate the methodology. After this, the single-dimension indicator results are combined into one index to represent the physical vQoSS.

(2) Subjective evaluation of perceived vQoSS is the score of willingness of street users to remain in the space, sampled from urban designers and borrowing widely accepted evaluation criteria from existing literature. This study uses enclosure, human-scale, transparency, tidiness and imageability as its standards and compares the physical vQoSS of streets with the perceived vQoSS of streets as a verification process. The perceived vQoSS (willingness to spend time in the space) is measured by a rating (score from 1 to 5) and taken as the benchmark to check the gap between automatic calculation of results and traditional subjective evaluation results.

(3) The identification of temporal variation of vQoSS of streets by detecting physical changes of streets follows. This study sections streets into four components, the façade, pedestrian space, motor road space, and the commercial space, or wall. The details of each part of the methodologies are as follows.

3.2.1. Physical vQoSS evaluation using SegNet and ArcGIS

SegNet and ArcGIS were chosen for the physical vQoSS measurement. According to the literature, physical vQoSS should consider enclosure, greenery, openness, scale, imageability, transparency and so on, but until now, the subtle factors such as tidiness, imageability, and transparency have been difficult to achieve through an auto-calculated

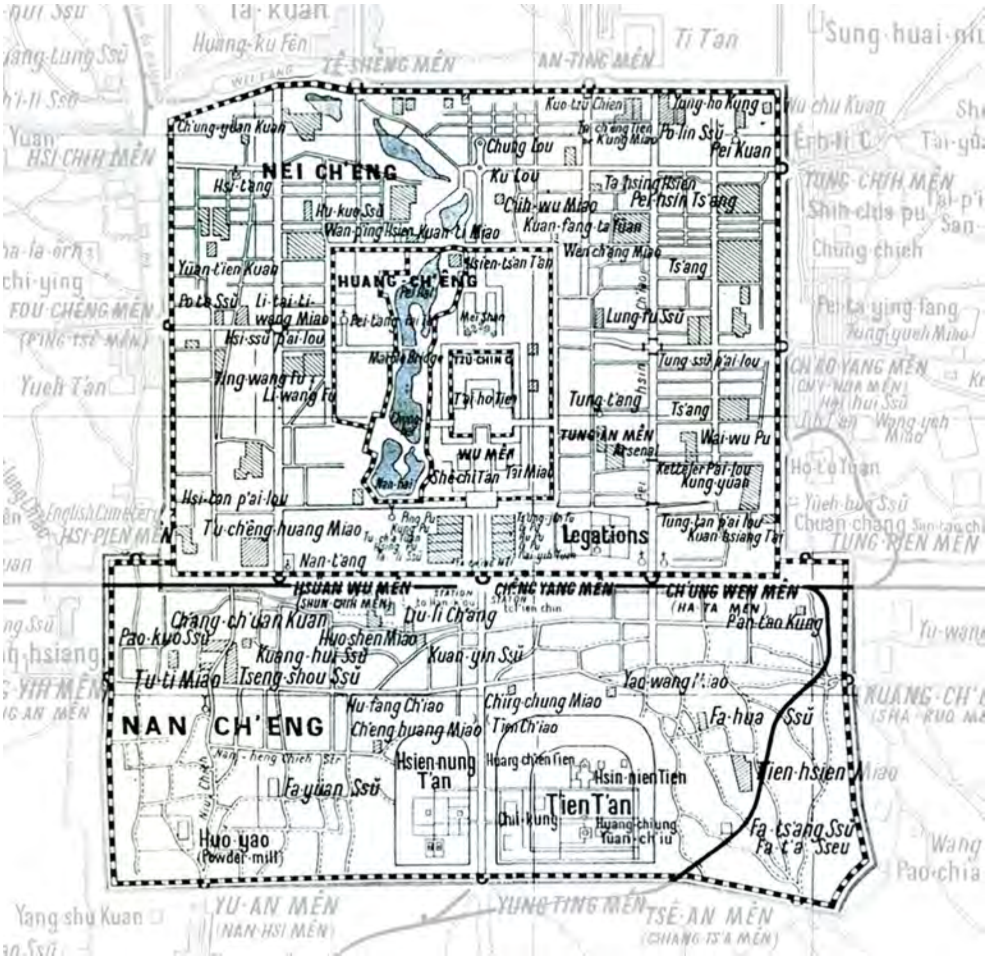


Fig. 2. *Peking and its Environs 1912* Map Source: Madrolle's Guide Books, http://www.lib.utexas.edu/maps/historical/history_china.html.

method. So rather than construct a concrete and all-inclusive index system, more effort is taken to construct an integrated method. Therefore, five important and measurable indicators are selected first: greenery, openness, and enclosure with the SegNet decoder (for the basic Algorithm principle, see: Kendall et al., 2015; Badrinarayanan et al., 2015), and then the cross-sectional proportion and the street wall continuity are measured with the newly emerging GIS approach provided by Harvey (2014). A lot of semantic pixel-width image segmentation methods based on convolutional networks have emerged recently, such as YOLO, ImageNet, SegNet, DeepLab and so on. SegNet is highly accurate and easily accessible. Fig. 5 is a segmentation demo by SegNet from their website (http://mi.eng.cam.ac.uk/projects/segnet/).

SegNet is a fully trainable approach for joint feature learning and image mapping to pixel-wise semantic labels, which turns out to be the first deep learning method to attach low resolution features with semantic labels (Kendall et al., 2015; Badrinarayanan et al., 2015). The convolutional neural network architecture of SegNet uses a set of encoders and decoders followed by a pixel-wise classifier as a core trainable engine (Badrinarayanan, et al., 2016), where the object recognition is achieved by large-sample labeling and training instead of color extraction. For indicators like greenery, openness, and enclosure, which are features that depend on the street composition ratio, SegNet could simplify the calculation approach and automate the interpretation of the streetscape. For example, for the recognition of "tree", the training label combines form and other features of trees, reducing the variability of seasons on the mapping. The following test samples of SegNet segmentation prove that tree form was ascertained correctly, even in leafless seasons and dusk scenes (Fig. 6). After testing a variety of outdoor RGB images, SegNet has been considered competitive in its measurement of the proportion of street elements, especially for buildings, column-poles, trees and the sky, and it has been used as a benchmark for the following improved segmentation method (Chen et al., 2016; Kundu, Vineet, & Koltun, 2016).

All of the most recent year's (2016) SVPs are interpreted into color groups using the SegNet decoder. Twelve kinds of physical components are obtained, namely sky, buildings, columns-poles, road markings, roads, pavement, trees, sign symbols, fences, vehicles, pedestrians, and bikes. We summarize percentages of each component in the four-directional (North, South, East, West) dataset (7568 images for 1892 locations) are summarized and the mean value for each location is calculated to represent the average condition of that location. Then the percentage of tree as greenery (%), sky ratio as openness (%), and the total sum of buildings, columns and trees as enclosure (%) are used. The total calculation is based on the resulting proportions calculated by SegNet.

In addition, methods are borrowed from Harvey (2014) using the ArcGIS geo-processing function to calculate cross-sectional proportion and street wall continuity. The method is a creative way to combine GIS with SVP data, automatically obtaining these two key physical indicators. Specifically, cross-sectional proportion is the ratio of average height to street width, which could be calculated directly for both sides. The height is the average height of the buildings. When there are only walls instead of buildings, it calculates with the height of the walls alongside, depending on their actual objective existence. While for the street wall continuity, the street is firstly broken into segments, 40 single-sided centerline buffers are then created at 1 m intervals, calculating the ratio of the area of each buffer and the area without the buildings. The street edge is selected out with SQL when the ratio

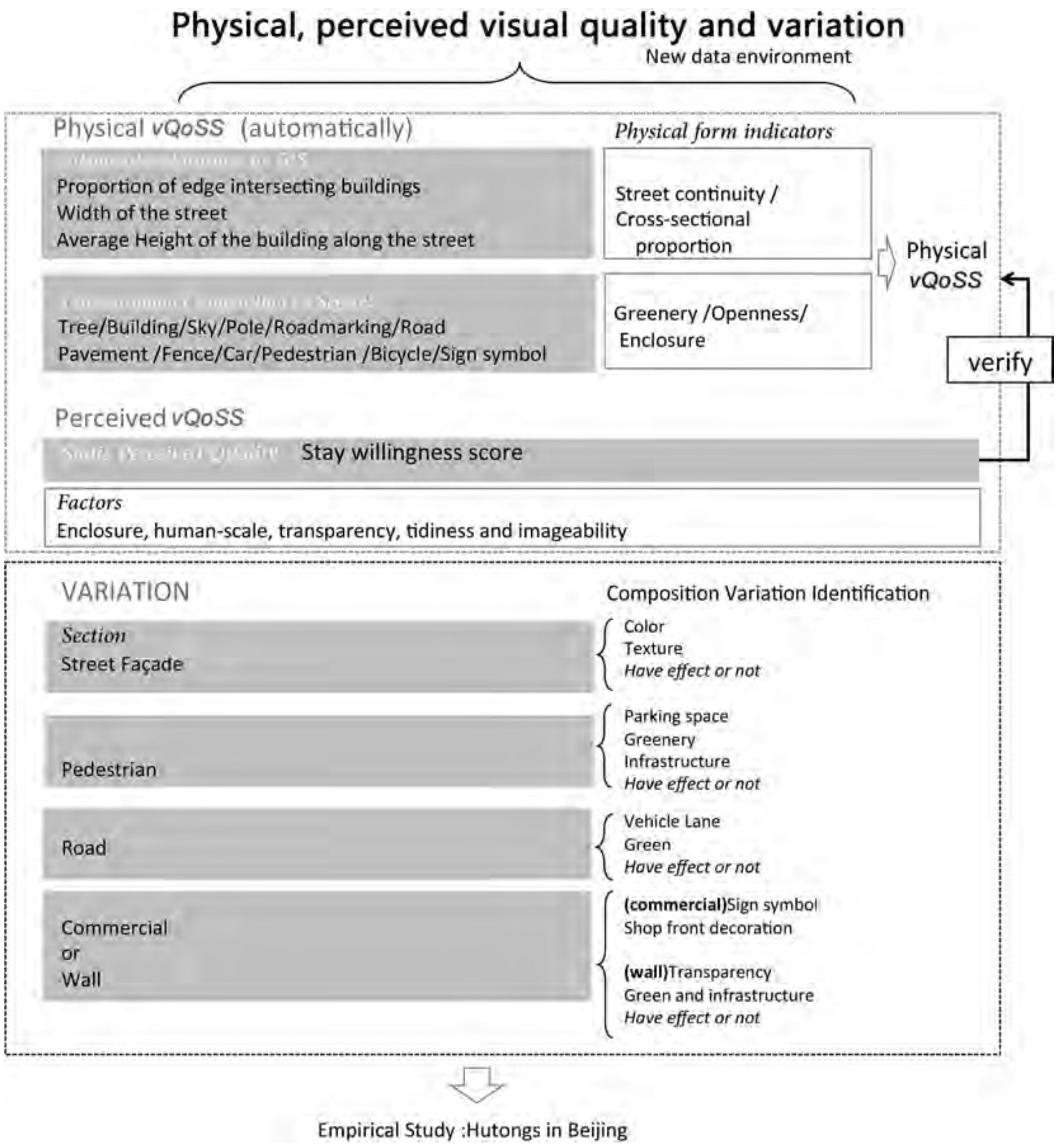


Fig. 3. The framework of the study, where 'Stay willingness score' is an individual's willingness to use or remain in a space.

reaches its peak. Continuity is the proportion of the length of the edge to the total length of the street segment. After this follows the calculation of the mean values of cross-sectional proportion and continuity of both sides as the value for that segment.

Thus, several indicators (a temporary five in this research) have been formed to comprehensively evaluate the physical vQoSS, the values of which a_i are normalized to the values z_i based on the deviation standardization Eq. (1):

$$z_i = \frac{a_i - \min(a_i)}{\max(a_i) - \min(a_i)} \tag{1}$$

i = greenery, openness, enclosure, wall continuity, cross sectional proportion and so on.

Following this, these five normalization values are summarized into one Index $b_{pquality}$ according to Eq. (2), below, which represents the overall physical vQoSS. Since the main purpose of this paper is to provide a methodology for quantitative measurement of visual quality of streetscapes in the empirical study of Hutongs the indicators are simplified into five and the coefficients of each are equally distributed for the convenience of discussion. Accurate coefficients still require further quantitative research based on large-sample surveying and theoretical literature reviews.

$$b_{pquality} = \sum_i \beta_i a_i \tag{2}$$

β_i are the weighting coefficients, awaiting further detailed research for improvement.

3.2.2. Evaluation of perceived vQoSS (Stay willingness scoring)

The willingness to stay (in a space) is scored by urban designers to measure the overall perceived vQoSS directly, borrowing criteria from Ewing and Clemente (2013). The five aspects represent the key aspects of perceived vQoSS, and the criteria have already been verified to successfully measure the previously unmeasured street quality. In order to clearly make comparison with the physical vQoSS result, they are summarized into one index, called the 'Stay willingness score'.

According to Ewing and Clemente (2013), *imageability* refers to the distinguishable characteristics of specific space; Lynch (1960) believes that space with good imageability is well shaped with special components for users to recognize, such as landmarks, walls, doorknobs, or domes, by which the street space gains a unique style; for the score of imageability, the raters value whether the space generates a feeling of recognizability, symbolism, indication of, or a sense of place or perception of locality. The more features above, the higher the score of

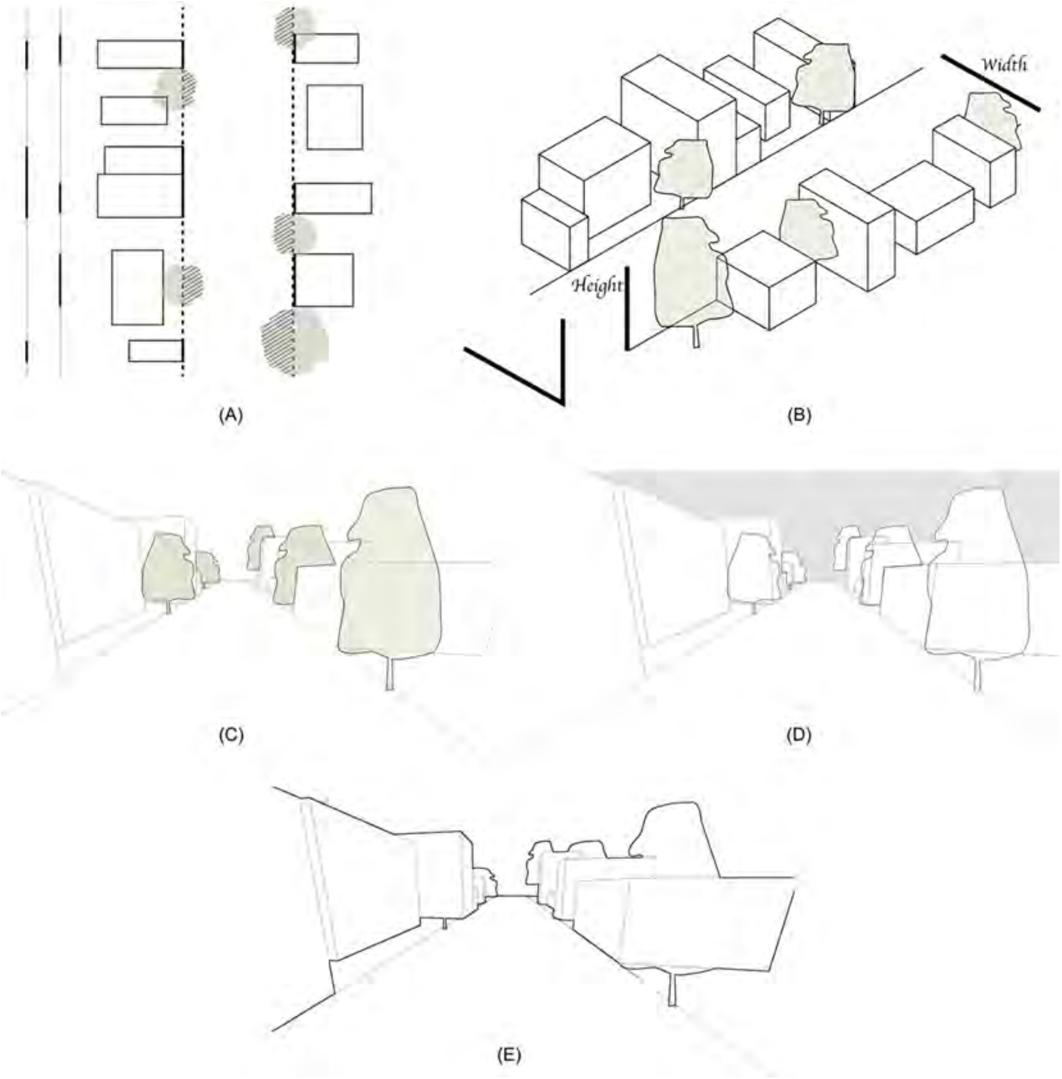


Fig. 4. Five selected aspects for physical vQoSS calculation in this paper. (A) Street wall continuity (B) Cross-sectional proportion (C) Greenery (D) Openness (E) Enclosure.

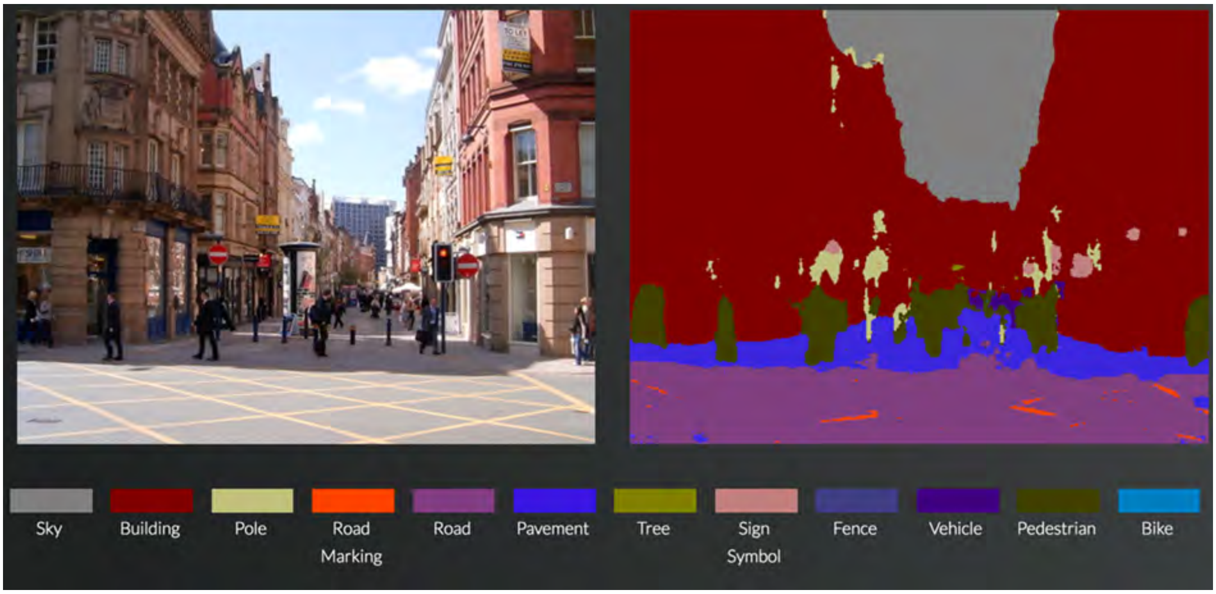


Fig. 5. A segmentation demo by SegNet.

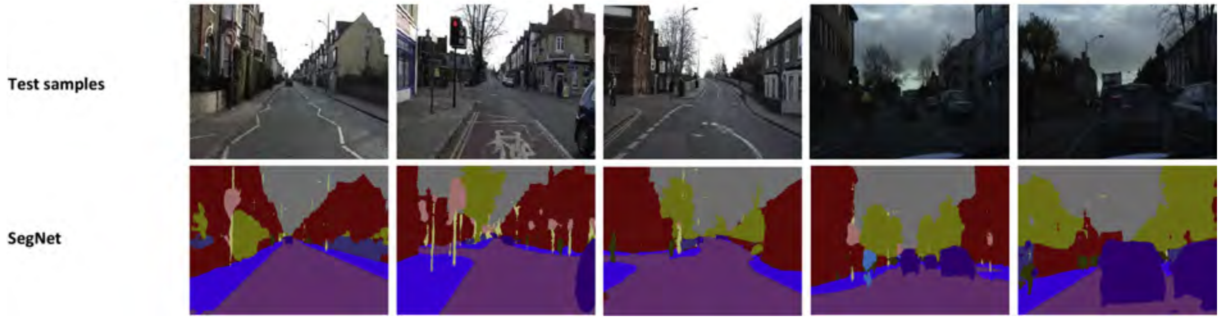


Fig. 6. test samples of SegNet source: Badrinarayanan, et al. (2016).

imageability. The meaning of *enclosure* has been explained in Section 2.1. *Human-scale*, which describes the scale and proportion of the physical street space, could affect psychological feelings through changes in architectural details, pavement forms, and greenery; *transparency* refers to the extent to which activities in the public street space could be visually accessed through transparent facades by the pedestrian, while *tidiness* is easy to understand, measuring the level of cleanliness or pollution and garbage.

Our four raters are all designers with a background in urban design education, and they were invited to comprehensively measure the vQoSS from the above five aspects, giving a score from 1 to 5. To avoid results from being affected by seasonal weather and subjective bias, they firstly scored together to have a consensus of the criteria, and then scored separately, which lasted nine days in total. Fig. 7 shows the sample SVPs for different Stay willingness scores.

3.2.3. *Variation recognition*

Variation recognition aims to reflect the appearance and changes in visual quality through variation of the street elements indirectly. Using current technology, the variation recognition is typically difficult to automate. Although multi-year SVPs for one location are initially intended to record the change for single sites, the recording angle could not avoid subtle deviation from year to year. As such, the artificial identification is still the most reliable method, which, however, is yet biased due to subjective judging. Therefore, this research aims at artificial recognition of variation by dividing the whole street section into four parts (Fig. 8) to minimize the resulting errors. From these four parts, the changing objects are identified in each area, then judged on whether those variations are effective for vQoSS improvement or not. As it shows in the evaluation decision tree (Fig. 9), each part has two or three elements to be checked (score 0 or 1) and an effective or non-effective indicator. If the variations have positive influence on vQoSS, then the score is 2, otherwise the score is 1 (No change is 0). Together, all these 11 sub-indicators make up a framework for observation. SVPs in Fig. 10 are the scoring samples for variation recognition.

4. Results

4.1. *Visual quality of Hutongs: Physical features*

Fig 11 shows the final five-aspect physical vQoSS index result and Fig. 12 shows the five physical features' results separately. The average physical quality index score is 1.39. The results are verified by comparing high scoring areas with site visits: it was found that the index could tell the physical form, but perceived vQoSS of (5) and (8) are low, though they are equipped with good form due to tidiness conditions. This means the index is accurate for physical form, but not for reflecting the comprehensive visual quality. It can be concluded that Hutongs on the north side of Chang'an Road have better skeletons than those on the south, especially Hutongs in the historical protection areas; those near the traditional imperial gardens (like *Houhai, Shishahai, Xihai,* and *Qianhai*), palaces, temples, traditional communities (*Fangjia Hutong, Dongsi Hutong Area*), and regeneration areas (*Dashilan Area, Nanluoguxiang*) received the highest score. It is not difficult to observe that some regeneration projects take full consideration of the historical forms (Fig. 13 (a)), keeping the original architectural features, while some partially change in spatial scale due to the need for modern transportation (Fig. 13(b)). Hutongs close to the main roads have a higher degree of openness and a relatively low degree of enclosure, indicating dramatic changes in the skeleton and widening processes catering to the demands of modern vehicles. In addition, in these Hutongs, the percentages of road is much higher than that of others. The cross-sectional proportion is generally lower than 0.5, while the continuity is around 0.7. Joint distributions of these two indicators show

Fig. 7. SVP samples for Stay willingness score ranging from 1 to 5.

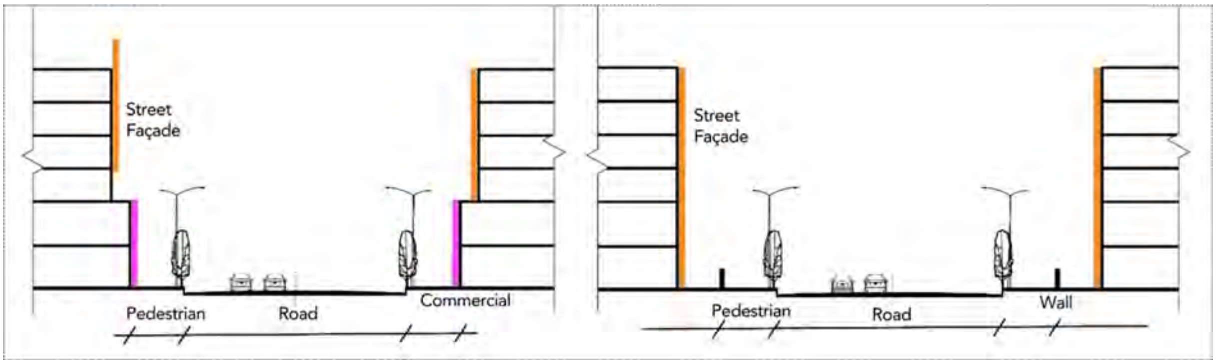


Fig. 8. Section of the street with commercial part along the Hutong (left) and without the commercial part, but with wall along the Hutong (right) Pink: Commercial space frontage; Orange: Street façade. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that less than 50% of Hutongs keep their compact skeleton, while others are replaced by porous forms.

In order to further portray the character of Hutongs and reflect their visual quality, an additional comparison is provided, calculating the physical component ratios of streets in Shanghai's Hengfu historical area, Beijing, and the Shanghai central area using the same method. The Hengfu historical area was previously the Shanghai French Concession, which retains western style streets and architecture. (Data was lacking for the other two aspects, otherwise all five aspects would have been compared together in the national context.) Fig. 14 shows the results for these four regions. From the comparison, it was found that most Hutongs are lacking visual greenery; the average percentage of greenery (19.8%) is similar to the central area of Beijing, but lower than that of the other three areas. Enclosures were much lower at 46.1%, because most Hutongs are 1-floor residential buildings with less tree shade and occlusions. These numeric results in the form illustrate a

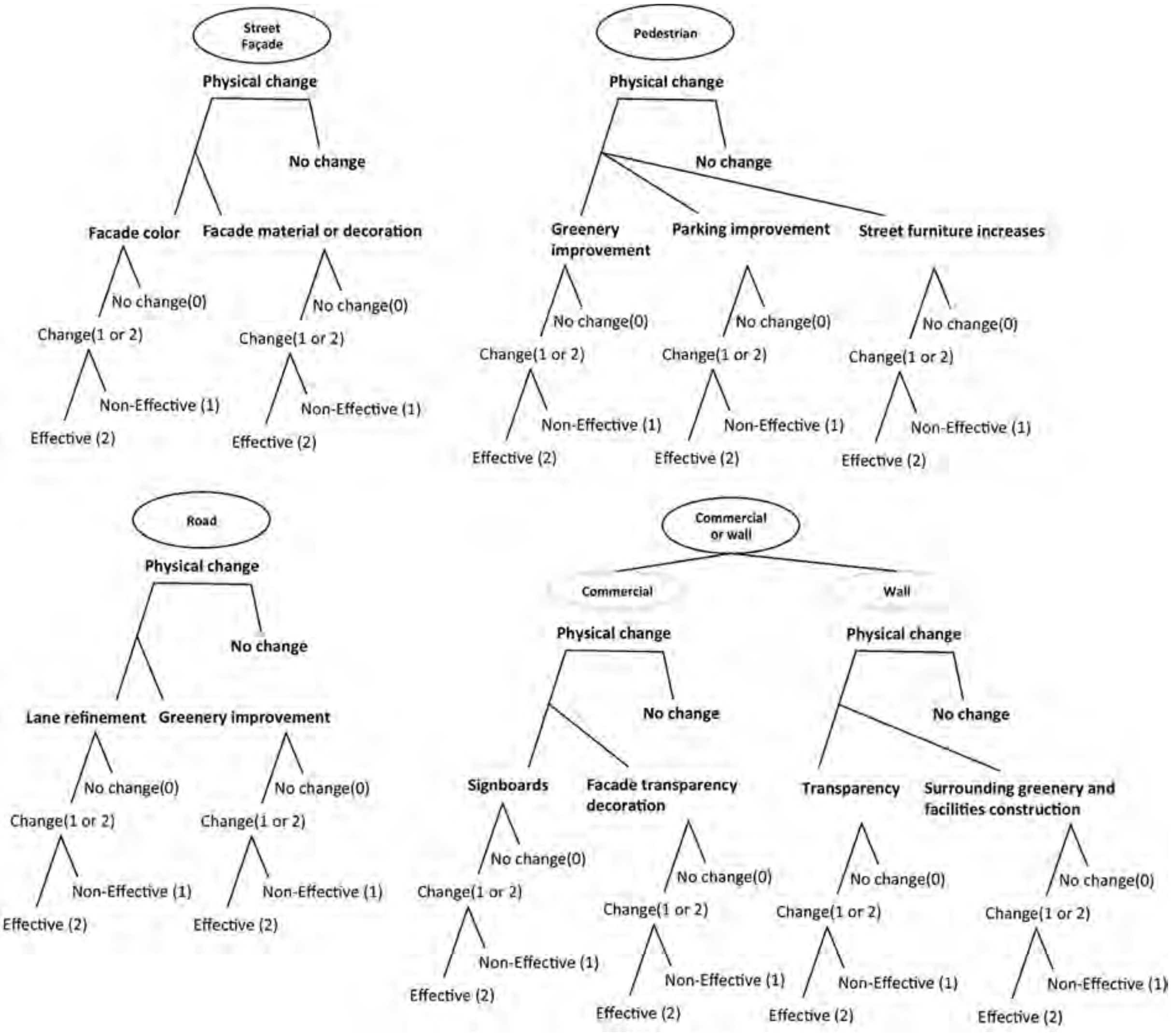


Fig. 9. Criteria for variation recognition.

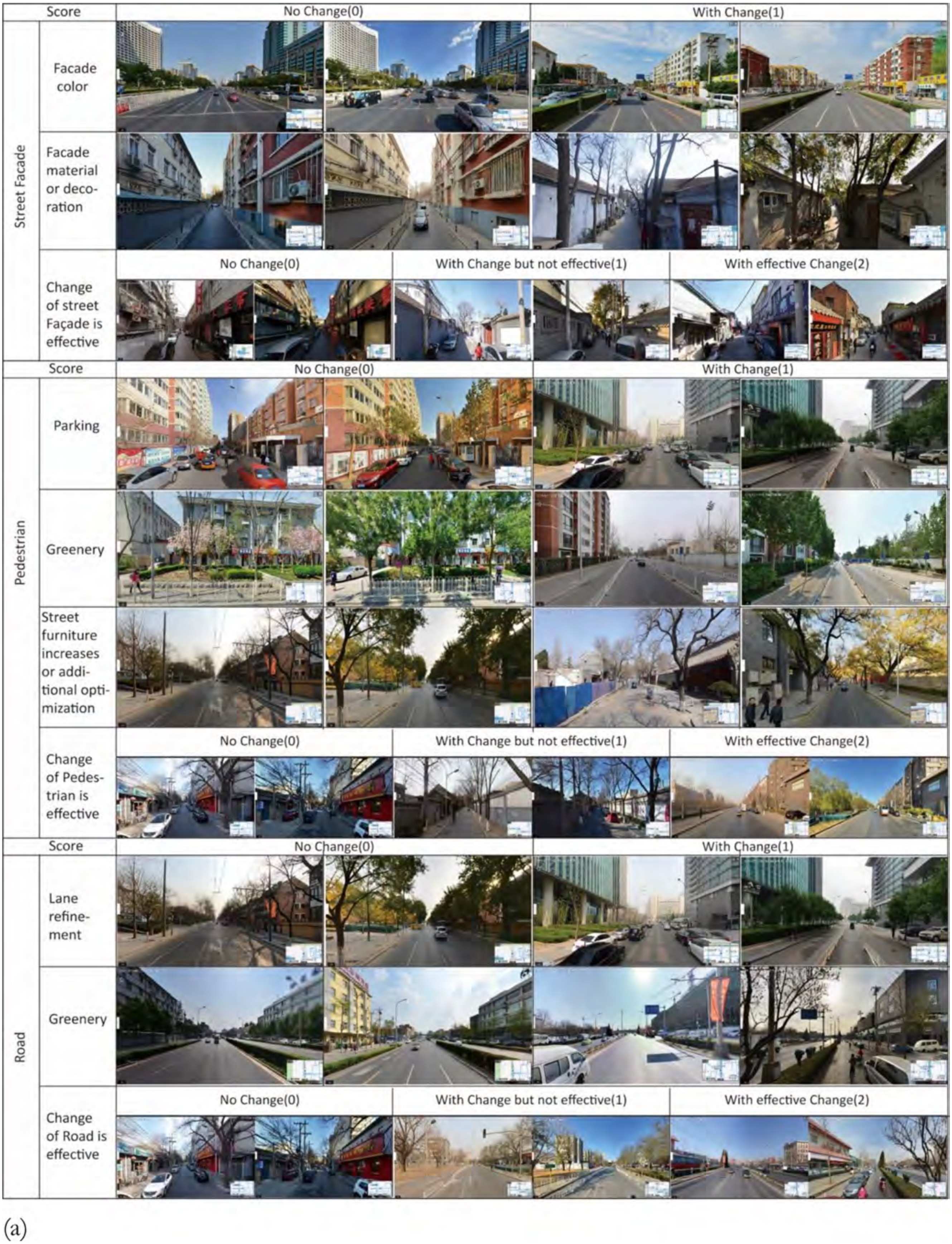


Fig. 10. Samples for four part variation recognition.

Fig. 10. (*continued*)

mid-level quality of Hutongs, which were lower than expectation.

4.2. Visual quality of Hutongs: Perceived stay willingness

Fig. 15 shows the overall Stay willingness scores of Hutongs. The average score is 2.59, where only 1.2% are rated 5, and 8.8% are rated 1. Most Hutongs are rated 2 or 3, accounting for 39.3% and 39.8% respectively. The result shows us the subjective understanding of Hutongs, that the overall perceived vQoSS is not one of satisfaction, except in some few areas (*Shichahai, Nanluoguxiang*) boasting improved Hutongs and roads (*Ganmian Hutong, Lumicang Hutong, Xigong Road, Middle Jingji Road*), which received higher ratings.

High score area in Fig. 14: (1) Northeast to the Shishahai, (2) Maoer Hutong and Nanluoguxiang, (3) Xisi Area, (4) Ganmian Hutong and Lumicang Hutong, (5) Middle Jingji Road, (6) Xigong West and East Road

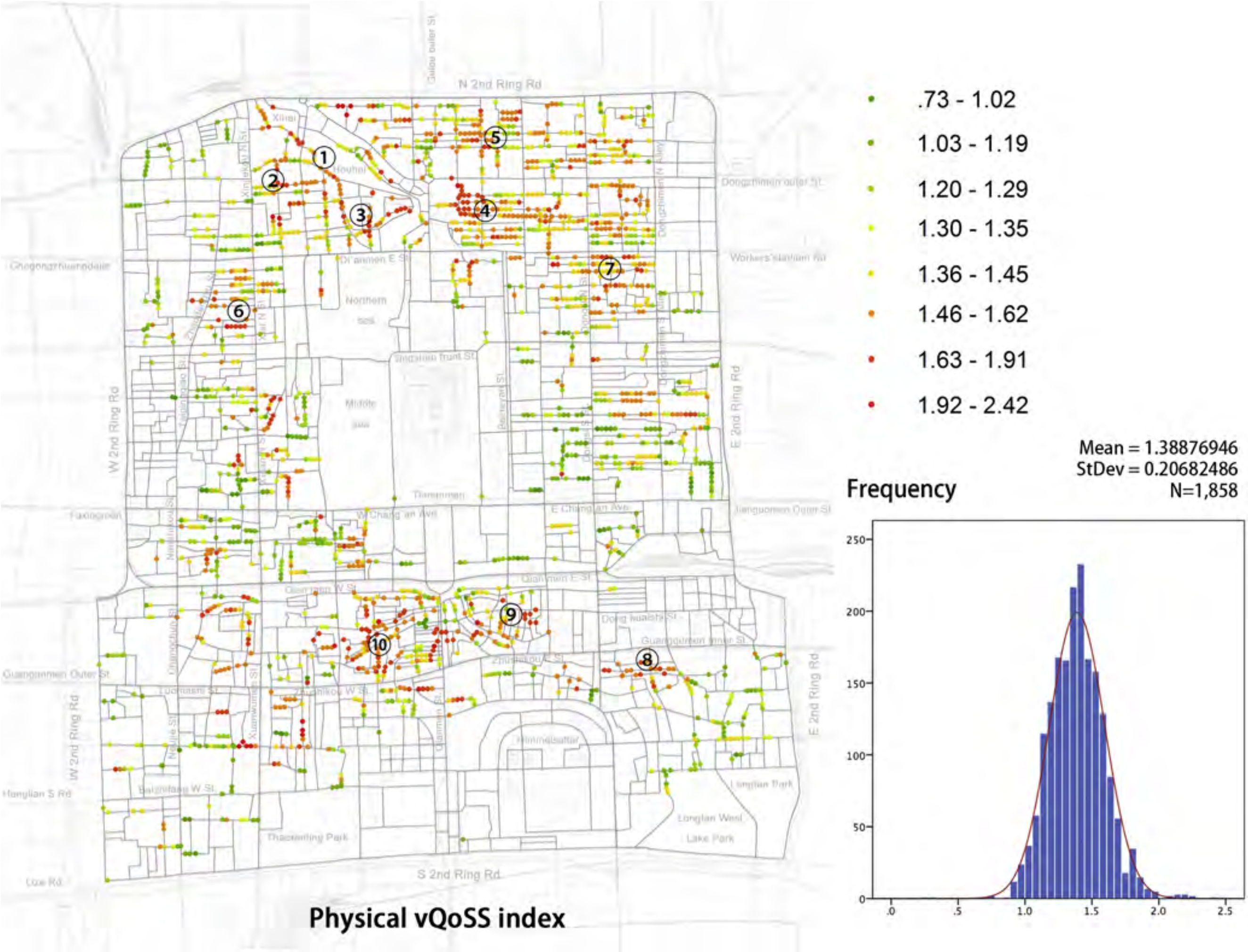


Fig. 11. Physical vQoSS Index of Hutongs. High scoring areas: (1) Eastern to Xihai, (2) North to Xinjiekou Street, (3) South to Gungwangfu, (4) Nanluoguxiang, (5) Fangjia Hutong, (6) North to Xisi Station, (7) Dongsi Hutong Area, (8) Xiting and Dongting Area, (9) Xixinglong Street, (10) Dashilan Area.

These results enabled the verification of the result of previous physical features calculated by automatic measurement. Taking the perceived vQoSS (the Stay willingness score) rated by educated designers as the benchmark, there is a discrepancy between the *D*-value of normalized perceived and physical vQoSS. Areas with inconsistent results are identified, which are shown as blue in Fig. 16, by the Standard Deviation classification, vQoSS of these locations is further verified by SVPs and site investigations. Results show that the main contributing factor to the difference between the approaches is the level of ignorance of aspects related to the degree of decline and human activity, such as the level of dilapidation on the street, tidiness, level of chaos, and vitality in the SVP auto-calculation process; while these factors are important indicators for subjective evaluation, they are missing from automated evaluation, leaving a gap for further research.

4.3. Visual quality of Hutongs: Temporal variation

Table 3 gives detailed statistical results for variation identification. Generally speaking, there has been very little variation in the elements over the past 3–4 years, mostly less than 10%. The total amount of variation varies from street section to section. Most changes take place in street façade and commercial real estate, accounting for 10% (Street Façade, Façade color), 9% (Street Façade, Façade material or decoration), 16% (Commercial, Signboards change) and 9% (Commercial, Store façade transparency, decoration change) respectively (see Table 3, below), possibly resulting from residential beautification and rapid store retail turnovers and replacements. 8% of the walls along the street changed transparency, 7% had changes with greenery and facilities construction, which could have promoted communication with the street. Only 284 Hutongs experienced variation regarding pedestrians, and 165 had road changes, accounting for around 5% variation. It is easy to conclude that vQoSS changes slowly and infrequently.

In addition, most of the variations are not satisfactory. The most effective improvements occur in the commercial section, driven by the market (1.7%), and the street façade (2.5%). Other sections account for less than 1.5% of effective improvements in total (Table 4). The renewal process for Hutongs seems to deliver unfavorable outputs, and it seldom reflects the concepts and principles of smart design, so more efforts are needed for urban regeneration with respect to the surrounding historical context. In other words, the improvement of the visual quality of Hutongs has been and will be a lengthy process.

5. Conclusion and discussion

5.1. Concluding remarks

This study proposed an evaluation framework for vQoSS, including physical and perceived vQoSS and assessment of its temporal variation. The first was calculated automatically with the support of SegNet, while the latter two were achieved quantitatively through a rating system based on systematic indicators. The realization of this combined method exploits new applications of Street View Picture and new

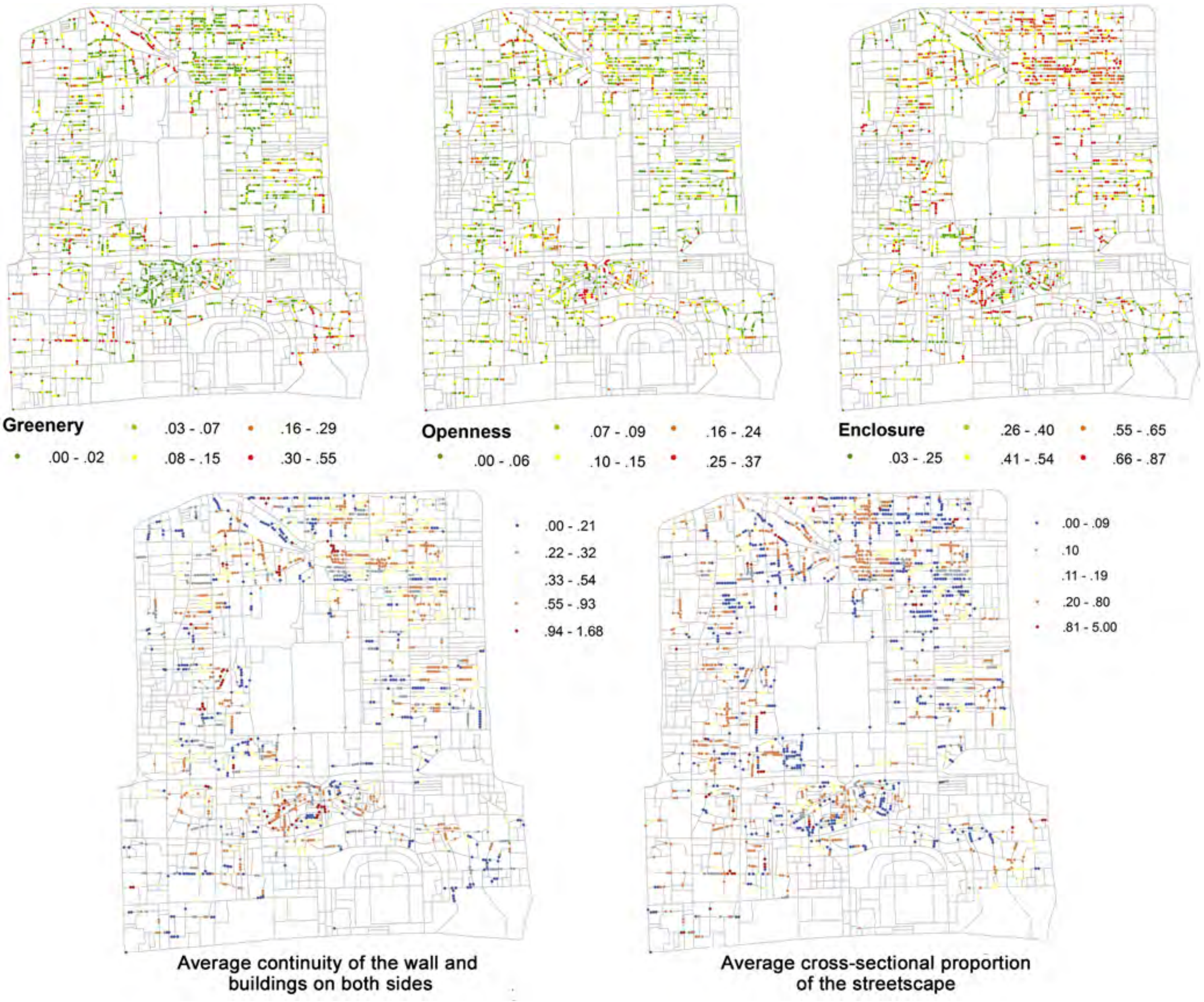


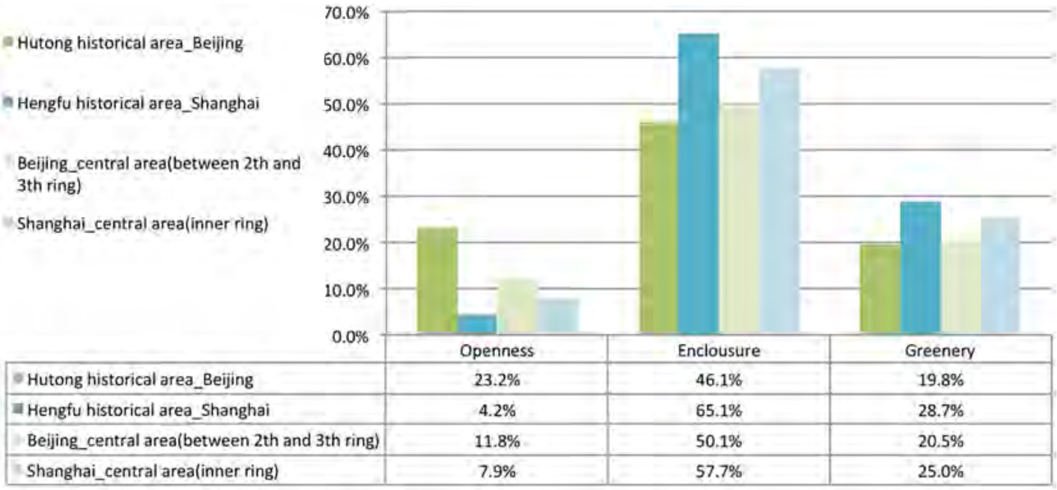
Fig. 12. Physical features of Hutongs.

emerging image segmentation techniques. With perceived vQoSS as the benchmark, the paper concludes the strengths and weaknesses of the automatic calculated method of physical vQoSS.

An empirical study of Hutongs shows that, although the Old City of Beijing has gained a reputation for its cultural, historical and architectural design values, the general visual quality fails to satisfy the current needs, remaining to be ameliorated. Physical feature analysis indicates Hutongs in the historical protection areas have better physical form, while others receive lower scores. However, joint analysis of street continuity and cross-sectional proportion show that less than half of the Hutongs have maintained their previous forms and structure while others are reconstructed in a relatively modern approach, replaced by porous forms. Many traditional styles and features in the inner city are challenged by modernization. The inner city is still a locus attracting developer attention. The process of improvement is an ongoing balance between the traditional and modern. Nevertheless, better skeletons do not mean better visual quality, –the degree of decline and human activity should be taken into consideration for the overall quality.

The variation results indicate that visual quality change is unobvious and slow; less than 10% of Hutongs have changing elements, while no more than 2.5% are effective in the past 3–4 years, indicating

Fig. 13. (a) left, Dashilan West street (regeneration) (b) right, Nanchizi Street.



	Openness	Enclousure	Greenery
Hutong historical area_Beijing	23.2%	46.1%	19.8%
Hengfu historical area_Shanghai	4.2%	65.1%	28.7%
Beijing_central area(between 2th and 3th ring)	11.8%	50.1%	20.5%
Shanghai_central area(inner ring)	7.9%	57.7%	25.0%

Fig. 14. Comparative analysis of greenery, openness, and enclosures between Hutongs in Beijing, the Shanghai Hengfu historical area, and the Beijing and Shanghai central areas. Footnote: the central areas in the two comparable largest cities, Beijing and Shanghai in China, were selected because their central areas remain relatively traditional after long-term urban development, of which the spatial scale is more comfortable and has been better maintained. The regions outside the central area are mostly newly developed high-rise blocks with wide roads. Therefore, it makes no sense to compare Hutongs to them. Secondly, to be specific, the Shanghai Hengfu historical area was picked as a reference because this area used to be a French Concession with a compact scale, pleasant greenery and good imageability.

the difficulties in renovation and regeneration of historical areas involved in chaotic property rights and interest-balancing issues. According to the site survey conducted later, the inner city is mixed with a variety of functions and property rights, varying from hostels, regenerated commercial streets, and low rent housing to residential houses. Therefore, to give Hutongs a new life and obtain a high-quality environment, public participation and environmental awareness are quite necessary. In addition, despite the less distinctive trend in environmental improvement, they show no sign of decline.

5.2. Potential applications

This study is an initial attempt to construct a quantitative research methodology for vQoSS. After testing, this methodology shows its feasibility and convenience for preliminary investigation of urban design, suggesting potential application as an augmented tool for designers to eliminate subjective judgment bias and criticism, and for enabling a more rational design. Considering the convenience and accessibility of data acquisition, it could be applied beforehand in evaluation, design elements extraction, and analysis of design implementation, so as to facilitate a feedback process, for example, as an assist-tool for the

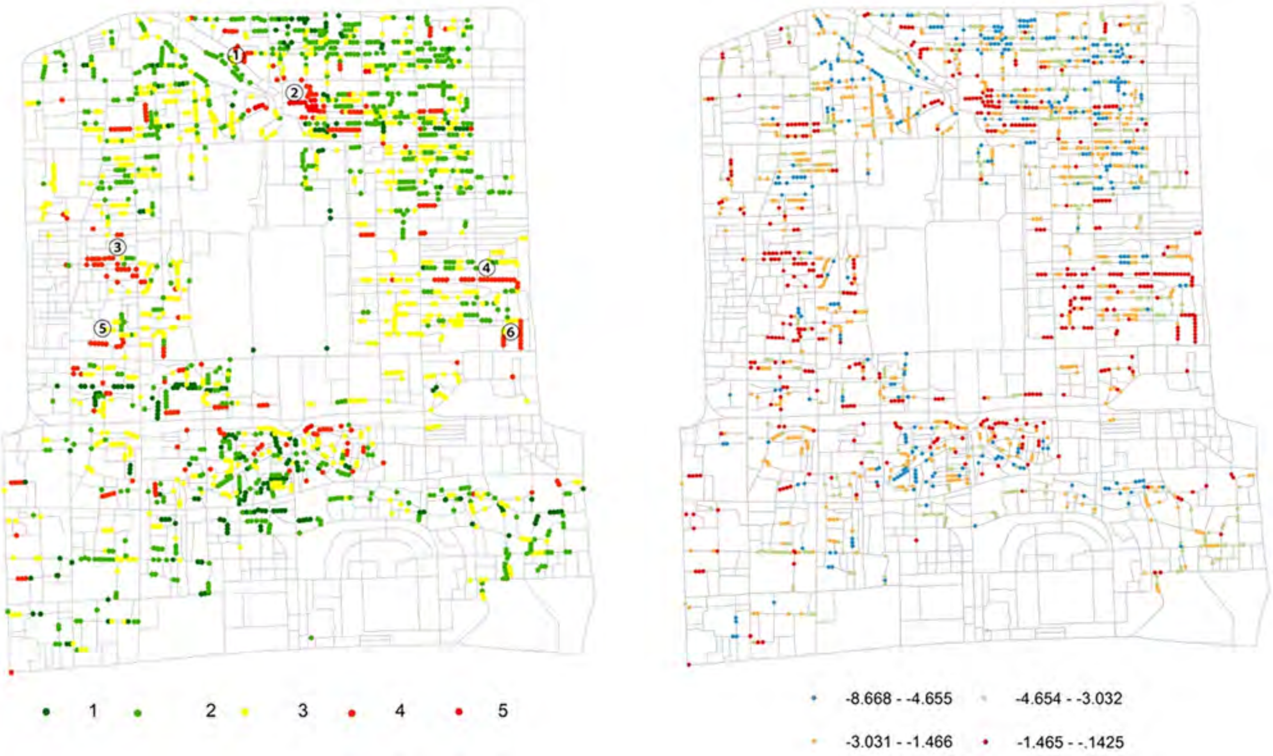


Fig. 15. Stay willingness scores of Hutongs.

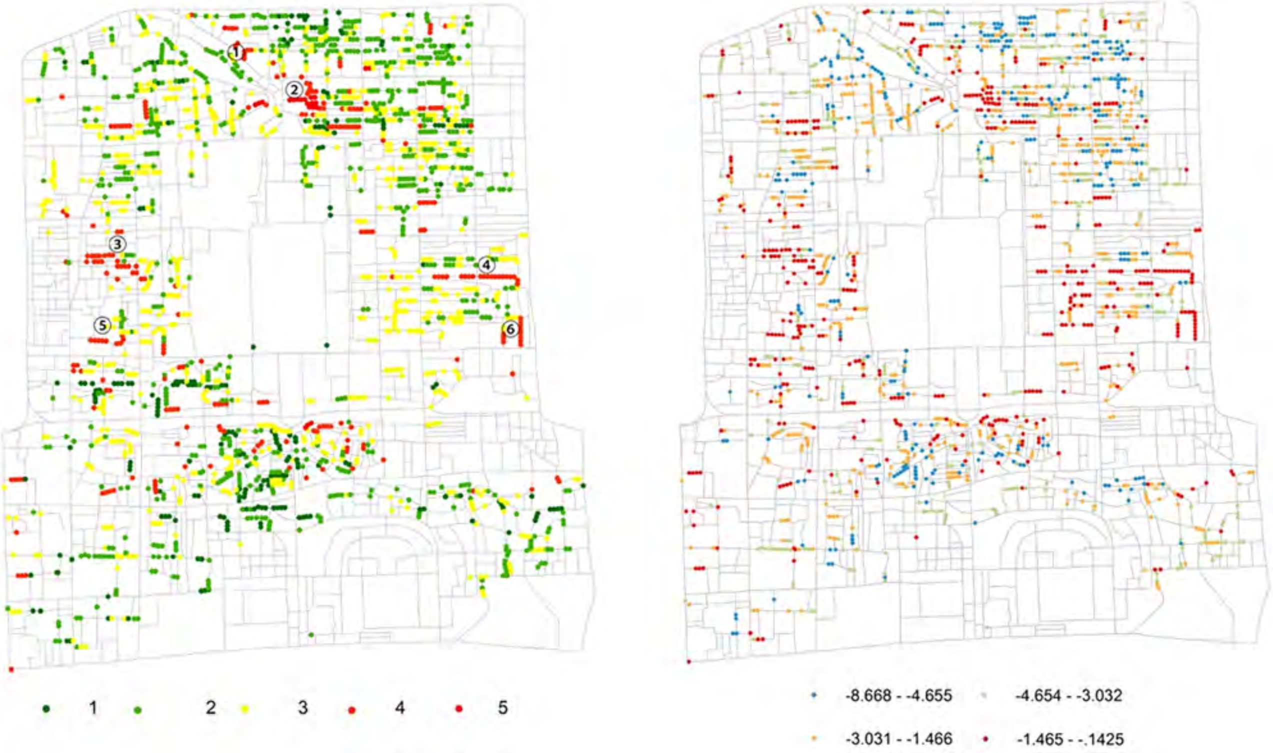


Fig. 16. D-value of perceived and physical vQoSS (Normalized by perceived vQoSS).

Table 3
Statistical result for variation.

	Indicators	N	Mean	Total change
Street Façade	facade color (X1)	1852	0.10	X1 + X2 = 352
	facade material or decoration (X2)	1852	0.09	
Pedestrian	Parking(X3)	1891	0.04	X3 + X4 + X5 = 284
	Greenery (X4)	1891	0.07	
	Street furniture increases or additional optimization (X5)	1891	0.04	
Road	Lane refinement (X6)	1498	0.06	X6 + X7 = 165
	Greenery improvement (X7)	1497	0.05	
Commercial	Signboards change (X8)	845	0.16	X8 + X9 = 211
	Store facade transparency, decoration change (X9)	842	0.09	
Wall	Transparency change (X10)	433	0.08	X10 + X11 = 68
	Surrounding greenery and facilities construction (X11)	441	0.07	

Table 4
Effectiveness of variation.

	The Effectiveness of Change	N	Proportion of the effective change
Street Façade	Non-Effective	7	3.5%
	Effective	16	2.5%
Pedestrian	Non-Effective	82	4.6%
	Effective	9	0.5%
Road	Non-Effective	56	3.9%
	Effective	10	0.7%
Commercial	Non-Effective	47	5.6%
	Effective	14	1.7%
Wall	Non-Effective	19	4.4%
	Effective	6	1.4%

assessment of master plans and comprehensive urban design. Designers could discover the key aspects and precise positions of street space for refinement, so that they could optimize the design and realize the optimal combination of science, practicality and aesthetics. Segmentation and objective physical component analysis aim to contribute empirical values for smart design guideline making in different function areas, such as residential, historical, business areas, and so on. In addition, the visual quality evaluation process also shows its potential use in stimulating participative planning. With a variety of evaluation platforms, street space quality can be rated directly by the residents, promoting users to express their expectations for the environment and forming a positive interaction between the citizens and the designers, truly enabling “streets for the people”.

5.3. Academic contributions

Beside the possibility for applying this method in urban design projects, this study also contributes a method for measuring the built environment as follows. First, the Street View Picture is proven to be a large-scale database, matching the need for visual quality data of the built environment. With the continuing development of Image Processing Technology, SVP may bring further findings in this field. Visual quality is one of the most important dimensions of the built environment. Using micro-scale SVP data, this study achieves auto-calculated physical vQoSS based on GIS and segmentation technology, and identifies the limitations by mutual verification and rectification with perceived vQoSS. Based on classical theories, a methodological framework is constructed for quantitative measurements of the streetscape, which is available for wide application. The partially inconsistent result provides clues for the potential for subsequent studies of detailed coefficients. Second, by extending the scale of the measureable built environment, the pattern of quality could be measured on a large scale with fine-granularity data (Fig. 17) from the previously

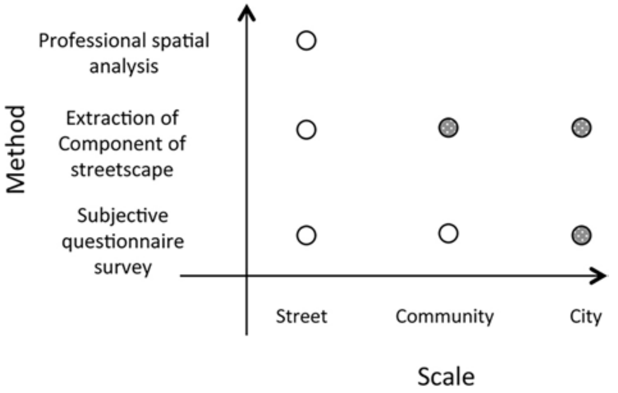


Fig. 17. Measurable scale expansion achieved by combining subjective and objective evaluation (in grey).

restricted community-scale to the current scale (including street, community and city) which comprehensively demonstrates the physical and social attributes of street space. Thus, the visual images of a city could be reflected through stitching streetscape images rather than only through qualitative description, while the large-scale calculation result also lays a foundation for potential investigation into the relation between the micro-built environment to macro-scale socioeconomic attributes in the city region. Third, this study provides an interdisciplinary perspective on built environment studies. Computer technology is foreseen to bring a qualitative leap through interdisciplinary integration of urban design and computing technology. Semantic image segmentation methods, like SegNet's, applied in this research, demonstrate a new approach for visual deconstruction, which may hopefully promote more automatic methods for smart design in the future.

5.4. Potential limitations and future research

While appreciating the merits of this study, there are still several limitations deserving further research in the near future. First, the concept of visual quality is human-dependent (subjective), considering its nature, based in sensation and perception. The scoring index for visual quality of street space may vary among different people groups, and future studies are expected to address this issue. Firstly, despite plenty of relevant theories and existing methods reviewed at the beginning of the article to ensure rationality and effectiveness, the selection of integrated indicators still suffers from the limitations of quantitative techniques. Second, the determination of coefficients (β_i) are still waiting for further comprehensive discussion, for which the streetscape typology based on function, location and hierarchy of streets are fundamental. Streetscape with more recreational function is supposed to give more weight to aesthetic indicators while for main streets focusing on mobility is on the opposite; In addition, the perceived quality of raters/users‘ judgments may be significantly affected by their expectation for vQoSS, therefore, the typology classification before data collection and rating is worth for cautious discussion. And for variation recognition, long-term field tracking surveys and corresponding time-serial SVP data are expected to be collected simultaneously, as a mutual verification of the computer-augment image recognition by traditional methods. Third, it was found that the physical vQoSS could only partially explain the overall vQoSS due to the ignorance of the degree of dilapidation, tidiness, chaos, and human vitality in this study, and the contribution of each indicator was thereby made unclear, requiring one-by-one detection with large sample data; special attention is required to bridge this gap. The best method for overcoming the limitation of perceived senses in the future, using intelligent technology to integrate the visual and non-visual factors, and providing an accurate guideline for city design implementation, is a topic needy of continuous discussion.

Acknowledgements

This work was supported by the National Water Pollution Control and Treatment Science and Technology Major Project of China (Grant No. 2017ZX07103-002), the National Nature Science Foundation of China (Grant No.51778319), and Tsinghua University (School of Architecture) CIFI Group Joint Research Center for Sustainable Community (Grant No.R201).

References

Aiello, L. M., Schifanella, R., Quercia, D., & Aletta, F. (2016). Chatty Maps: Constructing sound maps of urban areas from social media data. *Royal Society Open Science, 3*, 150690. https://doi.org/10.1098/rsos.150690.

Alcantarilla, P.F., Stent, S., Ros, G., Arroyo, R., & Gherardi, R., (2016). Streetview change detection with deconvolutional networks. In: Robotics: Science and systems conference (RSS).

Aletta, F., Lepore, F., Kostara-Konstantinou, E., Kang, J., & Astolfi, A. (2016). An experimental study on the influence of soundscapes on people's behaviour in an open public space. *Applied Sciences, 6*(10), 276.

Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). *A pattern language: Towns, buildings, construction.* New York: Oxford University Press.

Appleyard, D., & Lintell, M. (1972). The environmental quality of city streets: The residents' viewpoint. *Journal of the American Institute of Planners, 38*(2), 84–101.

Arnold, H. F. (1980). Trees in urban design. Trees in urban design.

Aspinall, P., Mavros, P., Coyne, R., & Roe, J. (2013). The urban brain: Analysing outdoor physical activity with mobile. *EEG. British Journal of Sports Medicine* bjsports-2012.

Bader, M. D., Mooney, S. J., Lee, Y. J., Sheehan, D., Neckerman, K. M., Rundle, A. G., & Teitler, J. O. (2015). Development and deployment of the Computer Assisted Neighborhood Visual Assessment System (CANVAS) to measure health-related neighborhood conditions. *Health & Place, 31*, 163–172.

Badrinarayanan, V., Handa, A., & Cipolla, R. (2015). Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. arXiv preprint arXiv:1505.07293.

Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J. F. (2009). Measuring the built environment for physical activity: State of the science. *American journal of preventive medicine, 36*(4), S99–S123.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual perception: Physiology, psychology, & ecology.* Psychology Press.

Carozza, L., Tingdahl, D., Bosché, F., & Gool, L. (2014). Markerless vision-based augmented reality for urban planning. *Computer-Aided Civil and Infrastructure Engineering, 29*(1), 2–17.

Cassidy, T. (1997). *Environmental psychology.* London, UK: Psychology Press.

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2016). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. arXiv preprint arXiv:1606.00915.

Curtis, J. W., Curtis, A., Mapes, J., Szell, A. B., & Cinderich, A. (2013). Using google street view for systematic observation of the built environment: Analysis of spatio-temporal instability of imagery dates. *International Journal of Health Geographics, 12*(1), 53.

Dover, V., & Massengale, J. (2013). *Street design: The secret to great cities and towns.* John Wiley & Sons.

Dupont, L., Antrop, M., & Van Eetvelde, V. (2014). Eye-tracking analysis in landscape perception research: Influence of photograph properties and landscape characteristics. *Landscape Research, 39*(4), 417–432.

Easteal, M., Bannister, S., Kang, J., Aletta, F., Lavia, L., Witchel, H.J. Urban sound planning in Brighton and Hove. In Proceedings of the Forum Acusticum, Krakow, Poland, 7–12 September 2014.

Ewing, R., & Clemente, O. (2013). *Measuring urban design: Metrics for livable places.* Washington, DC: Island Press.

Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban Design, 14*(1), 65–84.

Fruin, J. J. (1971). *Pedestrian planning and design.* New York: Metropolitan Association of Urban Designers and Environmental Planners.

Fyfe, N. (Ed.). (2006). *Images of the street: Planning, identity and control in public space.* Routledge.

Gehl, J. (2013). *Cities for people.* Washington, DC: Island press.

Gehl, J., & Gemzøe, L. (1996). *Public spaces-public life.* Copenhagen: Danish architectural press.

Gehl, J., & Svarre, B. (2013). *How to study public life.* Washington, DC: Island Press.

Grimmond, C. S. B., Potter, S. K., Zutter, H. N., & Souch, C. (2001). Rapid methods to estimate sky view factors applied to urban areas. *International Journal of Climatology, 21*(7), 903–913.

Harvey, C., Aultman-Hall, L., Hurley, S. E., & Troy, A. (2015). Effects of skeletal streetscape design on perceived safety. *Landscape and Urban Planning, 142*, 18–28.

Harvey, C. W. (2014). Measuring Streetscape Design for Livability Using Spatial Data and Methods.

Heath, G. W., Brownson, R. C., Kruger, J., Miles, R., Powell, K. E., & Ramsey, L. T. (2006). The effectiveness of urban design and land use and transport policies and practices to increase physical activity: A systematic review. *Journal of Physical Activity and Health,* *3*(s1), S55–S76.

Henshaw, V. (2014). *Urban smellscapes – Understanding and designing city smell environments.* Abingdon-on-Thames, UK: Routledge.

Jacobs, J. (1961). *The death and life of great American cities.* New York: Random House.

Jiang, B., Chang, C. Y., & Sullivan, W. C. (2014). A dose of nature: Tree cover, stress reduction, and gender differences. *Landscape and Urban Planning, 132*, 26–36.

Jiang, B., Larsen, L., Deal, B., & Sullivan, W. C. (2015). A dose–response curve describing the relationship between tree cover density and landscape preference. *Landscape and Urban Planning, 139*, 16–25.

Johansson, M., Sternudd, C., & Ferreira, I. (2015). The Walkshop: A tool to integrate research on human aspects of sustainable urban design in teaching. *Högre utbildning, 5*(3), 145–157.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology, 15*, 169–182.

Kaplan, R., & Kaplan, S. (1989). *The Experience of nature: A psychological perspective.* Cambridge: Cambridge University Press.

Kendal, D., Williams, N. S., & Williams, K. J. (2012). Drivers of diversity and tree cover in gardens, parks and streetscapes in an Australian city. *Urban Forestry & Urban Greening, 11*(3), 257–265.

Kendall, A., Badrinarayanan, V., & Cipolla, R. (2015). Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding. arXiv preprint arXiv:1511.02680.

Kundu, A., Vineet, V., & Koltun, V. (2016). Feature space optimization for semantic video segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3168–3175).

Li, X., Ratti, C., & Seiferling, I. (2017). Mapping urban landscapes along streets using google street view. *International cartographic conference* (pp. 341–356). Cham: Springer.

Li, X., Zhang, C., & Li, W. (2015). Does the visibility of greenery increase perceived safety in urban areas? Evidence from the Place Pulse 1.0 Dataset. *ISPRS International Journal of Geo-Information, 4*(3), 1166–1183.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening, 14*(3), 675–685.

Lin, T. P., Tsai, K. T., Hwang, R. L., & Matzarakis, A. (2012). Quantification of the effect of thermal indices and sky view factor on park attendance. *Landscape and Urban Planning, 107*(2), 137–146.

Long, Y., Liu, L., & Tang, J. (2016). How green are streets? An analysis for 245 major Chinese cities using Tencent street view. *One PLoS* under review.

Luigi, M., Massimiliano, M., Aniello, P., Gennaro, R., & Virginia, P. R. (2015). On the Validity of Immersive Virtual Reality as tool for multisensory evaluation of urban spaces. *Energy Procedia, 78*, 471–476.

Lynch, K. (1960). *The image of the city.* Cambridge: The MIT Press.

Lynch, K. (1984). *Good city form.* Cambridge: The MIT press.

McGinn, A. P., Evenson, K. R., Herring, A. H., Huston, S. L., & Rodriguez, D. A. (2007). Exploring associations between physical activity and perceived and objective measures of the built environment. *Journal of Urban Health, 84*(2), 162–184.

Middleton, J. (2016). The socialities of everyday urban walking and the 'right to the city'. Urban Studies, 0042098016649325.

Montgomery, C. (2013). *Happy city: Transforming our lives through urban design.* Macmillan.

Mounir, G. M., Attia, M. S., & Tayel, K. Y. (2016). Street children in Alexandria: Profile and psychological disorders. *Journal of High Institute of Public Health, 37*(1), 56–77.

Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L., & Hidalgo, C. A. (2015). *Do people shape cities, or do cities shape people? The co-evolution of physical, social, and economic change in five major US cities (No. w21620).* National Bureau of Economic Research.

Naik, N., Philipoom, J., Raskar, R., & Hidalgo, C. (2014). Streetscore–predicting the perceived safety of one million streetscapes. *2014 IEEE conference on computer vision and pattern recognition workshops* (pp. 793–799). IEEE.

Nasar, J. L. (1989). *Perception, cognition, and evaluation of urban places. In Public places and spaces.* Boston, MA: Springer31–56.

Owens, P. M. (1993). Neighborhood form and pedestrian life: Taking a closer look. *Landscape and Urban Planning, 26*(1–4), 115–135.

Pheasant, R. J., Horoshenkov, K., Watts, G., & Barret, B. T. (2008). The acoustic and visual factors influencing the construction of tranquil space in urban and rural environments tranquil spaces-quiet places? *Journal of the Acoustical Society of America, 123*(3), 1446–1457.

Rapoport, A. (2016). *Human aspects of urban form: Towards a man—environment approach to urban form and design.* Elsevier.

Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to audit neighborhood environments. *American Journal of Preventive Medicine, 40*(1), 94–100.

Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine, 25*(2), 80–91.

Sakurada, K., & Okatani, T. (2015). Change detection from a street image pair using CNN features and superpixel segmentation. In BMVC (pp. 61-1).

Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: Mapping the inequality of urban perception. *PLoS ONE, 8*(7), e68400.

Sallis, J. F., Johnson, M. F., Calfas, K. J., Caparosa, S., & Nichols, J. F. (1997). Assessing perceived physical environmental variables that may influence physical activity. *Research Quarterly for Exercise and Sport, 68*(4), 345–351.

Shen, Q., Zeng, W., Ye, Y., Arisona, S. M., Schubiger, S., Burkhard, R., & Qu, H. (2018). StreetVizor: Visual exploration of human-scale urban forms based on street views. *IEEE Transactions on Visualization and Computer Graphics, 24*(1), 1004–1013.

Slater, K. (1985). *Human comfort, Vol. 1.* Springfield, Ill., USA: CC Thomas.

Taneja, A., Ballan, L., & Pollefeys, M. (2011). Image based detection of geometric changes in urban environments. *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 2336–2343). IEEE.

Tang, J., & Long, Y. (2017). Evaluation for Quality of Street Space in center area of mega-city: A comparative analysis of Beijing and Shanghai. *Planners, 2017*(2), 68–73 (In Chinese).

Tang, J., Ma, Y., Zhai, W., & Long, Y. (2016). Measuring quality of street space, its temporal variation and impact factors: An analysis based on massive street view pictures. *New Architecture, 2016*(5), 130–135 (In Chinese).

Thompson, C. W. (2002). Urban open space in the 21st century. *Landscape and Urban Planning, 60*, 59–72.

Unger, J. (2008). Connection between urban heat island and sky view factor approximated by a software tool on a 3D urban database. *International Journal of Environment and Pollution, 36*(1–3), 59–80.

Wang, X., Kim, M. J., Love, P. E., & Kang, S.-C. (2013). Augmented Reality in built environment: Classification and implications for future research. *Automation in Construction, 32*, 1–13.

Ye, Y., & van Nes, A. (2014). Quantitative tools in urban morphology: Combining space syntax, spacematrix, and mixed-use index in a GIS framework. *Urban Morphology, 18*(2), 97–118.